Path-Phase Information Complementarity for Interfering Particles through State-Discrimination

Noam Erez,* Daniel Jacobs, and Gershon Kurizki
Department of Chemical Physics, Weizmann Institute of Science, Rehovot, 76100, Israel

We analyze the trade-off between the amounts of information obtainable on complementary properties of a qubit state by simultaneous measurements. We consider a "state discrimination" scenario wherein the same measurements are repeated, but the input states must be guessed in every run. We find a general complementarity relation for path-phase guesses by any generalized measurements in this scenario. The counterpart of this input-output mutual information (MI) reveals a hitherto unknown aspect of complementarity.

PACS numbers: 03.67.-a, 42.50.Dv, 42.50.Ex

arXiv:0903.4256v1 [quant-ph] 25 Mar 2009

As is well known, the measurement of one observable "disturbs" a complementary observable, i.e., introduces uncertainty in it. A complementarity or duality relation has been derived[1, 2]and experimentally verified[3]for Hilbert space of dimensionality 2. This relation quantifies path predictability versus fringe visibility of a particle in a Mach-Zehnder interferometer (MZI) with a partly efficient which-path detector (Fig. 1a).This relation reads:

$$D^2 + V^2 \leq 1 \tag{1}$$

The path distinguishability, D, is related to the which-way probability, $\mathcal{P}_{\rm ww}$, of guessing the path correctly for a *known* input state and a which-way detector of efficiency (reliability) $E \leq 1$, while the fringe visibility, V, is related to the which-phase probability, $\mathcal{P}_{\rm WP}$, of guessing correctly which MZI port the particle will exit through (for an optimal choice of the phase between the arms) [2, 4]

$$\mathcal{P}_{\rm WW} = \frac{1+D}{2}, \quad \mathcal{P}_{\rm WP} = \frac{1+V}{2}. \tag{2}$$

Yet, in this setup, the WW and WP probabilities refer to two *alternative* measurements[5]. Indeed, $\mathcal{P}_{\rm WP}$ is the probability of predicting correctly where the particle will exit (inset (a) of Fig. 1a).By contrast, $\mathcal{P}_{\rm WW}$ is operationally meaningful only in a measurement (inset (b) of the figure). 1) where the exit beam splitter of the MZI is removed, because only then can the readout of the partly efficient WW detector be verified. Thus, in the scheme of Fig. 1, simultaneous guesses of path and phase cannot be verified or falsified in the same predictive experiment, i.e., either $\mathcal{P}_{\rm WP}$ or $\mathcal{P}_{\rm WW}$ must represent a counterfactual probability. We may think of this "predictive" duality as a constraint on the optimal strategies in a single-player game, in which the player (Bob) knows the initial state and the experimental setup and tries to guess the outcome of each measurement.

Is it possible to obtain a duality relation for path and phase information such that both have *simultaneous* operational meaning in each experimental run? As we show,

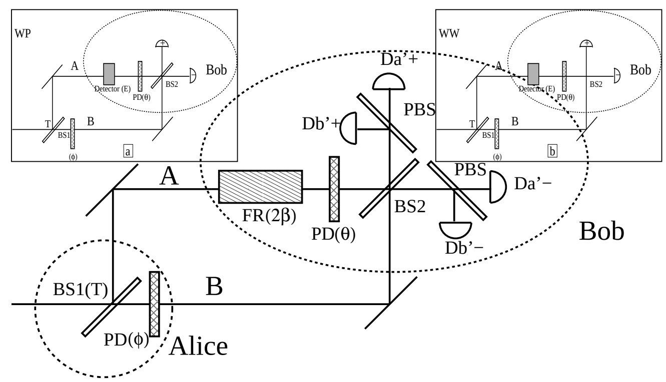


Figure 1: Setup for predictive complementarity game. The initial state is specified by a phase delay ϕ and transmissivity T at BS1. Bob can either measure the phase (inset a), or the path (inset b). Faraday rotator (FR) serves to correlate the path with the polarization.

such a relation can indeed be given in the context of quantum *state discrimination*, namely, measurements aimed at optimally guessing the initial state out of a set of possible states[6].In contrast to the "predictive" WW-WP duality, the proposed "retrodictive" duality described below is a bound on the optimal strategies in a two-player game: the guessing by Bob which of the several alternative input states had been prepared by Alice prior to the one measurement Bob performed. The precise rules of the game for this state-discrimination scenario are as follows:

(1) Alice randomly chooses to prepare the particle in one of the four input states (Fig. 2a):

$$|b_{\rm ww}, b_{\rm wp}\rangle_{\alpha,\phi} \equiv T(b_{\rm ww}\alpha)|A\rangle + e^{b_{\rm wp}i\phi}T(-b_{\rm ww}\alpha)|B\rangle \tag{3}$$

Here $|A, B\rangle$ are the path states (represented by qubit states $|\sigma_z = \pm 1\rangle$), $T(\pm b_{\rm ww}\alpha) = \cos\left(\frac{\pi}{4} \pm \frac{b_{\rm ww}\alpha}{2}\right)$ the corresponding amplitudes, and $e^{b_{\rm wp}i\phi}$ their relative phase factor. The four input states correspond to the choices of the parameters $(b_{\rm ww} = \pm 1, b_{\rm wp} = \pm 1)$.

(2) Bob receives the qubit, and after performing a mea-

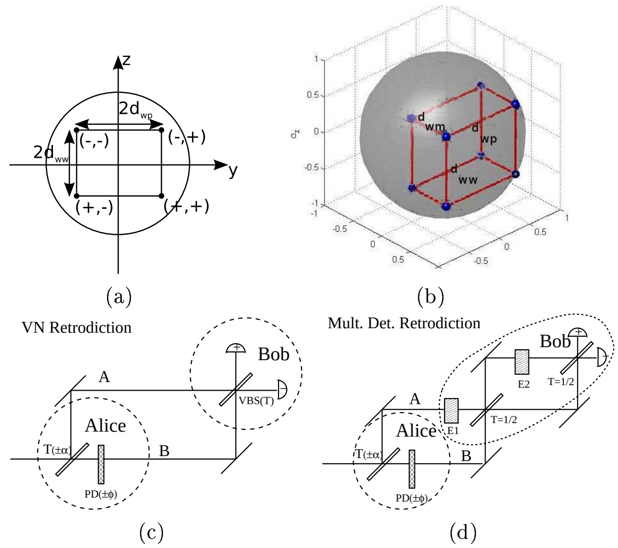


Figure 2: (a) Bloch representation of Alice's four alternative WW-WP input states (labeled by $b_{\rm ww}$ and $b_{\rm wp}$), with distances $d_{\rm ww}$ and $d_{\rm wp}$. (b) Bloch representation of 8 WW-WP-WM input states. (c) VN scheme (tunable output BS). (d) Multiple inefficient detector scheme for mixed state discrimination.

surement of his choice, tries to guess the values of the two bits $b_{\rm ww}, b_{\rm wp}$ (which are statistically independent), i.e., guess which of the four possible input states was chosen by Alice.

For a given choice of Alice's parameters (α, ϕ), each strategy that Bob adopts yields probabilities $P_{\rm WW}$ and $P_{\rm WP}$ to correctly guess $b_{\rm ww}$ and $b_{\rm wp}$, respectively (Roman font is henceforth used for $P_{\rm WW}$ and $P_{\rm WP}$, as opposed to the calligraphic font for "predictive" probabilities $\mathcal{P}_{\rm ww}$ and $\mathcal{P}_{\rm wp}$ above). Bob's strategy is "Pareto optimal"[7]if there is no other strategy that yields a pair of probabilities ($P_{\rm WW}$, $P_{\rm WP}$) such that one is strictly better and the other not worse than its counterpart. The set of all optimal pairs is called the "Pareto Frontier".

The optimal probability of distinguishing two states of a qubit with density operators ρ_1 and ρ_2 is given by 1/2 plus their trace distance[8]: $d_{\rm Trace}(\rho_1, \rho_2) \equiv \frac{1}{2}{\rm Tr}|\rho_1 - \rho_2|$. This is half the Euclidean distance between the corresponding Bloch vectors. Let us define $d_{\rm ww} \equiv d_{\rm Trace}(|b_{\rm ww} = +1, b_{\rm wp}\rangle, |b_{\rm ww} = -1, b_{\rm wp}\rangle), d_{\rm wp} \equiv d_{\rm Trace}(|b_{\rm ww}, b_{\rm wp} = +1\rangle, |b_{\rm ww}, b_{\rm wp} = -1\rangle)$. Here $d_{\rm ww}$ denotes the which-way distinguishability of the set of inputs, while $d_{\rm wp}$ is the which-phase distinguishability.

The Bloch vector corresponding to the input state $\mathbf{b}_i^{\rm in} = |b_{\rm ww}, b_{\rm wp}\rangle_{\alpha,\phi}$ in Eq. (3) is:

$$\mathbf{r}_i = d_0\hat{\mathbf{x}} + b_{\rm wp}d_{\rm wp}\hat{\mathbf{y}} + b_{\rm ww}d_{\rm ww}\hat{\mathbf{z}}, \tag{4}$$

where $d_0 = \cos\alpha\cos\phi$; $d_{\rm ww} = \sin\alpha$; $d_{\rm wp} = \cos\alpha\sin\phi$. The set of Alice's allowed input states form a rectangle on the Bloch Sphere with dimensions given by $2d_{\rm ww}$ and $2d_{\rm wp}$. (Fig. 2a).

Bob is allowed to perform generalized measurements (POVMs), as well as projective (Von Neumann) ones, by letting the particle interact with an ancillary system (ancilla) and then performing a projective measurement on both together (the WW detector in Fig. 1 is just such an ancilla). Any POVM can be represented by a set of operators, $\{A_i\}_{i=1\ldots N}$, satisfying[9] $A_i \geq 0, A_i^\dagger = A_i, \sum_i A_i = 1$. Operationally, this means that when performing the corresponding generalized measurement on a system initially in state ρ, the ith "measurement" outcome appears with probability ${\rm Tr}\{A_i\rho\}$.

Lemma A POVM $\{A_j\}$ on a qubit can be described as a collection of weighted points in the Bloch Ball: $A_j = \mu_j\frac{1+\mathbf{R}_j\cdot\boldsymbol{\sigma}}{2}$, with $\mu_j \geq 0, \|\mathbf{R}_j\| \leq 1$. Furthermore, every such POVM has a "refinement" consisting of weighted points on the Bloch *Sphere* ($\|\mathbf{R}_j\| = 1$), such that all information on the original POVM can be retrieved from it. Thus, without loss of generality, we shall assume a representation of the latter form. The proof of this lemma is given in the Supplement.

By assumption, all inputs $\mathbf{b}_\mathbf{i}^{in}$ are equally probable: $p_\mathbf{i} = \frac{1}{4}$. The joint input-output distribution is then: $p_{\mathbf{i},j} = p_\mathbf{i}Tr\{\rho_\mathbf{i}A_j\} = \frac{\mu_j}{8}(1+d_0x_j+i_{\rm wp}y_jd_{\rm wp}+i_{\rm ww}z_jd_{\rm ww})$, where $\mathbf{R}_j = (x_j, y_j, z_j)$.

Theorem .1 *For any POVM, the WW and WP probabilities satisfy:*

$$\left(\frac{2P_{\rm WW}-1}{d_{\rm ww}}\right)^2 + \left(\frac{2P_{\rm WP}-1}{d_{\rm wp}}\right)^2 \leq 1 \tag{5}$$

Equality holds iff all the Bloch vectors have the form $\mathbf{R}_j = (0, \pm\sqrt{1-z_0^2}, \pm z_0)$, *with* $z_0 \in [0,1]$, *and the corresponding weights* $\mu_{\pm,\pm}$ *satisfy:* $\mu_{++} = \mu_{--} = \mu, \ \ \mu_{+-} = \mu_{-+} = 1-\mu$ *for some* $\mu \in [0,1]$.

Outline of proof As shown in detail in the Supplement, the probability of correctly guessing $b_{\rm ww}$ is given by:

$$P_{\rm WW} = \sum_j \max\{p_{i_{\rm ww},j}, p_{-i_{\rm ww},j}\} = \frac{1}{2}\left(1+\sum\frac{\mu_j}{2}|z_j|\,d_{\rm ww}\right), \tag{6}$$

$$P_{\rm WP} = \frac{1}{2}\left(1+\sum\frac{\mu_j}{2}|y_j|\,d_{\rm wp}\right). \tag{7}$$

Combining these two equations, we have

$$\left(\frac{2P_{\rm WW}-1}{d_{\rm ww}}\right)^2+\left(\frac{2P_{\rm WP}-1}{d_{\rm wp}}\right)^2 \leq \sum\frac{\mu_j}{2}\left(z_j^2+y_j^2\right) \leq 1. \tag{8}$$

■

In what follows we shall restrict ourselves to measurements (POVMs) which are Pareto optimal, unless stated

otherwise. Two particular instances of this optimal class are of special interest. The case $\mu \in \{0,1\}$ describes a Von-Neumann measurement in the $x-z$ plane (this can be realized by choosing the phase delays of the output beam splitter of the MZI appropriately). The case $\mu = \frac{1}{2}$ describes the measurement with a WW detector, as in Fig. 1a.

For the WW-detector assisted measurement (Fig. 1a), the probabilities are determined by the detector efficiency, E: $z_0 = E$. Likewise, in the Von-Neumann measurement corresponding to an output BS with transmissivity $\frac{1}{2} \leq T_{\mathrm{Out}} \leq 1$ and correctly chosen input phases (Fig. 2c), $z_0 = 1 - 2T_{\mathrm{Out}}$.

The (unique) measurement minimizing the overall probability of error (in guessing both path and phase) is also Pareto optimal:

Theorem .2 *A POVM that maximizes the probability of guessing the path and phase simultaneously is necessarily WW-WP Pareto-optimal.*

Proof: A similar calculation to that used in the proof of Theorem .1 gives the probability of correctly guessing the input state, $P_c = \frac{1}{2}\left(P_{\mathrm{WW}} + P_{\mathrm{WP}} - \frac{1}{2}\right)$. Clearly, increasing one of the partial probabilities without reducing the other implies improving the average ■

WW- and WP-Information complementarity. These results can be recast in information-theoretic terms. Let $\mathbf{b}^{\mathrm{in}} = (b_{\mathrm{ww}}, b_{\mathrm{wp}})$ be the two random (statistically independent) bits Alice chooses for her *input* state, as before, and denote the observables measured by Bob by $\mathbf{b}^{\mathrm{out}}$. This set of *classical* stochastic variables (albeit related by a quantum channel) allows the definition of the WW(WP)- output information as the mutual information between $b_{\mathrm{ww(wp)}}$ and b^{out}: $I_{WW} \equiv I(b_{\mathrm{ww}} : \mathbf{b}^{\mathrm{out}})$; $I_{WP} \equiv I(b_{\mathrm{wp}} : \mathbf{b}^{\mathrm{out}})$.

To separate out the WW correlation explicitly, we introduce a new stochastic variable constructed out of the fundamental ones:

$$WW = \begin{cases} 1 & P(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}) > P(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}) \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where $P(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}})$ is the apriori probability for $b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}$ to take the values which actually occurred. We see from Eq. (9) that $P_{\mathrm{WW}} = P(WW = 1)$. This also implies $H(WW) = H_2(P_{\mathrm{WW}})$, where $H_2(x) = H(x, 1-x)$ is the binary entropy. We define an analogous variable WP, such that $P_{WP} = P(WP = 1)$.

Now, given the value of $\mathbf{b}^{\mathrm{out}}$, b_{ww} and WW determine each other, so that they are interchangeable, in the joint entropy $H(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}) = H(WW, \mathbf{b}^{\mathrm{out}})$. Since for optimal measurements, WW is stochastically independent of $\mathbf{b}^{\mathrm{out}}$, as can be seen from $P_{\mathrm{WW}} = \frac{1}{2}(1 + z_0 d_{\mathrm{ww}})$ (see Supplement), it follows that $H(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}) = H(WW, \mathbf{b}^{\mathrm{out}}) = H(WW) + H(\mathbf{b}^{\mathrm{out}})$. By definition, the mutual information $I(b_{\mathrm{ww}} : \mathbf{b}^{\mathrm{out}}) \equiv H(b_{\mathrm{ww}}) + H(\mathbf{b}^{\mathrm{out}}) - H(b_{\mathrm{ww}}, \mathbf{b}^{\mathrm{out}}) = H(b_{\mathrm{ww}}) - H(WW)$. From this last equation (and its WP analog) follow the intuitively appealing relations:

$$I_{\mathrm{WW}} = 1 - H_2(P_{\mathrm{WW}}), \quad I_{\mathrm{WP}} = 1 - H_2(P_{\mathrm{WP}}). \tag{10}$$

We note that I_{WW} and I_{WP} are monotonically increasing functions of P_{ww} and P_{wp}, respectively (because $P_{\mathrm{ww}}, P_{\mathrm{wp}} \in [\frac{1}{2}, 1]$). Thus, the P_{WW}-P_{WP} complementarity of Eq. (5) implies a similar trade-off for I_{WW} and I_{WP}. The amount of information about Alice's input settings contained in Bob's measurement results is given by the mutual information between them: $I_{\mathrm{in-out}} = I(\mathbf{b}^{\mathrm{in}} : \mathbf{b}^{\mathrm{out}})$. The relation between $I_{\mathrm{in-out}}$ and I_{WW}, I_{WP} is given by the following:

Theorem .3

$$I_{\mathrm{in-out}} = I_{\mathrm{WW}} + I_{\mathrm{WP}} + I_{\mathrm{WW:WP}}. \tag{11}$$

Proof: (1) For the important special case $\mu = \frac{1}{2}$ (WW-detector scheme), the output observables also consist of two bits $\mathbf{b}^{\mathrm{out}} = (b_{\mathrm{ww}}^{out}, b_{\mathrm{wp}}^{out})$, which turn out to be statistically independent. As the notation suggests, the WW output bit (obtained from the reading of the WW detector) has non-zero mutual information with b_{ww}, but not with b_{wp}, and conversely for the WP output bit. From

$$\begin{aligned} &I(b_{\mathrm{ww}}, b_{\mathrm{ww}}^{\mathrm{out}} : b_{\mathrm{wp}}, b_{\mathrm{wp}}^{\mathrm{out}}) \equiv \\ &H(b_{\mathrm{ww}}, b_{\mathrm{ww}}^{\mathrm{out}}) + H(b_{\mathrm{wp}}, b_{\mathrm{wp}}^{\mathrm{out}}) - H(b_{\mathrm{ww}}, b_{\mathrm{wp}}; b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) \end{aligned} \tag{12}$$

and

$$\begin{aligned} &H(b_{\mathrm{ww}}, b_{\mathrm{ww}}^{\mathrm{out}}) = H(WW) + 1, H(b_{\mathrm{wp}}, b_{\mathrm{wp}}^{\mathrm{out}}) = H(WP) + 1 \\ &H(b_{\mathrm{ww}}, b_{\mathrm{wp}}; b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) = H(WW, WP) + 2, \end{aligned} \tag{13}$$

we obtain:

$$\begin{aligned} &I(b_{\mathrm{ww}}, b_{\mathrm{ww}}^{\mathrm{out}} : b_{\mathrm{wp}}, b_{\mathrm{wp}}^{\mathrm{out}}) \equiv \\ &H(WW) + H(WP) - H(WW, WP) = I(WW : WP) \end{aligned} \tag{14}$$

Comparing

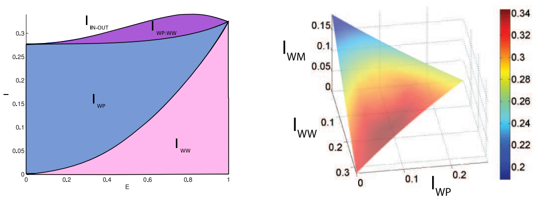
$$\begin{aligned} &I_{\mathrm{in-out}} = I(b_{\mathrm{ww}}, b_{\mathrm{wp}} : b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) \equiv \\ &H(b_{\mathrm{ww}}, b_{\mathrm{wp}}) + H(b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) - H(b_{\mathrm{ww}}, b_{\mathrm{wp}}; b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) \end{aligned} \tag{15}$$

with Eq. (12), and using the independence of $b_{\mathrm{ww}}, b_{\mathrm{wp}}$ and of $b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}$ to decompose: $H(b_{\mathrm{ww}}, b_{\mathrm{wp}}) = H(b_{\mathrm{ww}}) + H(b_{\mathrm{wp}})$, $H(b_{\mathrm{ww}}^{\mathrm{out}}, b_{\mathrm{wp}}^{\mathrm{out}}) = H(b_{\mathrm{ww}}^{\mathrm{out}}) + H(b_{\mathrm{wp}}^{\mathrm{out}})$, we get Eq. (11).

The third term, $I_{WW:WP}$, is a novel corollary of our treatment. This "cross-information"has its origin in the fact that, although b_{ww} is independent of b_{wp} and of b_{wp}^{out}, and b_{ww}^{out} is also independent of b_{wp} and of b_{wp}^{out}, b_{ww}^{out} and b_{ww} together tell us something about b_{wp}^{out} and b_{wp}. Namely, a correct (incorrect) WW guess implies lower (higher) probability of correct WP guess. If Bob is allowed to bet on WW first, and is then told the outcome,

this will not change the value of $b_{\rm wp}$ he bets on, but will affect the odds of his guessing correctly! This is a *hitherto unnoticed subtle form of complementarity.*

(2) In the general case ($\mu \in [0,1]$), $\mathbf{b}^{out}$ does not decompose neatly into WW and WP parts, and proceeding as above, one ends up with the relation: $I_{\rm in-out} = I_{\rm WW} + I_{\rm WW} +$ $[I(b_{\rm ww}, \mathbf{b}^{\rm out} : b_{\rm wp}, \mathbf{b}^{\rm out}) - H(\mathbf{b}^{\rm out})]$. The expression in the square brackets is known as the conditional mutual information: $I(b_{\rm ww} : b_{\rm wp}|\mathbf{b}^{\rm out})$ and is equal to $I_{\rm WW:WP}$, as required. ■



(a) (b)

Figure 3: (a) Illustration of (11) parametrized by the detector efficiency, E in a WW-detector scheme (input parameters: $d_{\rm ww} = .65, d_{\rm wp} = .6$). (b) Pareto optimal surface with color representing the magnitude of I_{in-out} (input parameters: $d_{\rm ww} = .65$, $d_{\rm wp} = .6$, $d_{\rm wm} = 0.5$).

Complementarity for mixed states The geometry of the trace-distance suggests the following generalization of the problem to 3 independent input parameters for mixed states. We now consider 8 input states $\rho_{b_{\rm ww}=\pm1, b_{\rm wp}=\pm1, b_{\rm wm}=\pm1}$ corresponding to the Bloch vectors:

$$\mathbf{r}_{b_{\rm ww}, b_{\rm wp}, b_{\rm wm}} = d_0\hat{\mathbf{x}} + b_{\rm wm} d_{\rm wm}\hat{\mathbf{x}} + b_{\rm wp} d_{\rm wp}\hat{\mathbf{y}} + b_{\rm ww} d_{\rm ww}\hat{\mathbf{z}}, \tag{16}$$

where $d_{\rm wm}$ is the mixing distance (a measure of impurity), defined similarly to $d_{\rm ww}$ and $d_{\rm wp}$, and the distances satisfy: $d_{\rm ww}^2 + d_{\rm wp}^2 + d_{\rm wm}^2 \le 1$ and $d_0 = \sqrt{1 - (d_{\rm ww}^2 + d_{\rm wp}^2)} - d_{\rm wm}$. The set of all eight states is now a rectangular box within the Bloch Ball (Fig. 2b). The Pareto frontier is now an ellipsoid:

$$\left(\frac{2P_{\rm WW} - 1}{d_{\rm ww}}\right)^2 + \left(\frac{2P_{\rm WP} - 1}{d_{\rm wp}}\right)^2 + \left(\frac{2P_{\rm WM} - 1}{d_{\rm wm}}\right)^2 = 1. \tag{17}$$

Schemes for path-phase-"mixedness" complementarity are shown in Fig. 2c and 2d (see supplement). The input-output mutual information for optimal schemes is:

$$I_{\rm in-out} = I_{\rm WW} + I_{\rm WP} + I_{\rm WM} + C_{\rm WW:WP:WM}. \tag{18}$$

Here $C_{\rm WW:WP:WM}$ is the "total correlation" $C(WW : WP : WM)$ defined by: $C(X,Y,Z) \equiv H(X) + H(Y) + H(Z) - H(X,Y,Z)$. It is the mixed-state counterpart of the novel cross-information term $I_{\rm WW:WP}$ in (11) (see discussion following (15)).

To conclude, while the "predictive" distinguishability-visibility duality holds for *two* alternative measurements (itself a manifestation of complementarity), our state-discrimination scenario allows for only one type of measurement: we have more preparations, instead. This scenario yields the generalized complementarity and mutual information (MI) relations (5) and (11), and their mixed input generalization, (17) and (18). These results show that complementarity may be reformulated as an information tradeoff obtained on complementary properties (path and phase) in a single measurement. The structure of this information tradeoff is richer than previously thought, as it allows for "cross information": information gained on the odds of WW guess given the WP result (or vice versa). Since the goal of this work is to allow for *simultaneous* WW and WP MI, our scenario is *not* the sequence-reversed version of standard predictive duality. However, the latter scenario is of great interest to quantum cryptography[10]. As shown in [11], one can derive Pareto-optima for the latter problem from those of the present one quite simply. Hence, the tight bound on MI in (11), (18) is expected to be useful for improving the corresponding one in the cryptographic setting.

We acknowledge the support of the EC, GIF and ISF.

* Electronic address: nerez@weizmann.ac.il

[1] J. Wheeler and W. Zurek, *Quantum theory and measurement* (Princeton University Press, New Jersey, 1982);W. K. Wootters and W. H. Zurek, Phys. Rev. D **19**, 473 (1979);D. M. Greenberger and A. Yasin, Physics Letters A **128**, 391 (1988);G. Jaeger, A. Shimony, and L. Vaidman, Phys. Rev. A **51**, 54 (1995).

[2] B.-G. Englert, Phys. Rev. Lett. **77**, 2154 (1996);B.-G. Englert and J. A. Bergou, Optics Communications **179**, 337 (2000).

[3] S. Dürr, T. Nonn, and G. Rempe, Nature **395**, 33 (1998);V. Jacques et al.Phys. Rev. Lett. **100**, 220402 (2008).

[4] M. Kolár, T. Opatrný, N. Bar-Gill, N. Erez, and G. Kurizki, New Journal of Physics **9**, 129 (2007);M. Kolář, T. Opatrný, and G. Kurizki, Optics Letters **33**, 67 (2008).

[5] A. Luis, Phys. Rev. A **70**, 062107 (2004).

[6] U. Herzog and J. A. Bergou, Phys. Rev. A **65**, 050305 (2002);Phys. Rev. A **70**, 022302 (2004).

[7] A. Colell, M. Whinston, and J. Green, *Microeconomic Theory* (1995);P. Straffin, *Game Theory and Strategy* (The Mathematical Association of America, 1993).

[8] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, (2000).

[9] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Pub, 1995).

[10] S. Wu, S. Yu, and K. Mølmer, Phys. Rev. A **79**, 022320 (2009).

[11] N. Erez, D. Jacobs, and G. Kurizki (2009), arXiv/0903.1921.

Supplement

Lemma A POVM $\{A_j\}$ on a TLS, can be described as a collection of weighted points in the Bloch Ball: $A_j = \mu_j \frac{1+\mathbf{R}_j\cdot\boldsymbol{\sigma}}{2}$, with $\mu_j \geq 0, \|\mathbf{R}_j\| \leq 1$. Furthermore, every such POVM has a "refinement" consisting of weighted points on the Bloch *Sphere* ($\|\mathbf{R}_j\| = 1$), such that all information on the original POVM can be retrieved from it.

Proof: Any hermitian operator A on a qubit's Hilbert space is a linear combination of the Pauli matrices and the identity: $A = \frac{\mu}{2}1 + \mathbf{R}\cdot\boldsymbol{\sigma}$. If A is further required to be positive, then we must have $\|\mathbf{R}\| \leq \frac{\mu}{2}$, as can be seen from the requirement $\left\langle \hat{\mathbf{R}}\cdot\sigma = -1 \middle| A \middle| \hat{\mathbf{R}}\cdot\sigma = -1 \right\rangle \geq 0$. Hence $A_i = \mu_i \frac{1+\mathbf{R}_i\cdot\boldsymbol{\sigma}}{2}$ with $\mu_i \geq 0, \|\mathbf{R}_i\| \leq 1$, which is, by definition, in the Bloch ball. The condition that the operators sum to the identity implies $\frac{1}{2}\sum \mu_i = 1; \sum \mu_i \mathbf{r}_i = 0$. Also notice that every POVM has a "refinement" consisting of weighted points on the Bloch Sphere, such that all information on the original POVM can be retrieved from its refinement. The refinement is obtained by replacing each operator A_i whose Bloch vector lies inside the ball $((\mu_i, \mathbf{R}_i), |\mathbf{R}_i| < 1)$, by the pair $\{A_i^{\pm}\}$, where $A_i^{\pm} \leftrightarrow \left(\mu_i \frac{1\pm|\mathbf{R}_i|}{2}, \pm\hat{\mathbf{R}}_i\right)$.

■

Theorem .1 For any POVM, the WW and WP probabilities satisfy:

$$\left(\frac{2P_{\mathrm{WW}}-1}{d_{\mathrm{ww}}}\right)^2 + \left(\frac{2P_{\mathrm{WP}}-1}{d_{\mathrm{wp}}}\right)^2 \leq 1 \qquad (19)$$

Equality holds iff the Bloch vectors all have the form $\mathbf{R}_j = (0, \pm\sqrt{1-z_0^2}, \pm z_0)$, with the same $z_0 \in [0,1]$, and the corresponding weights $\mu_{\pm,\pm}$ satisfy: $\mu_{++} = \mu_{--} = \mu$, $\mu_{+-} = \mu_{-+} = 1-\mu$ for some $\mu \in [0,1]$.

Proof: Let $\mathbf{i} = 1\ldots 4$ stand for each possible $(i_{\mathrm{ww}}, i_{\mathrm{wp}})$ pair, as described in Eq. (3). By assumption all inputs $\mathbf{i}$ are equally probable, and we use the notation: $p_{\mathbf{i}} = \frac{1}{4}$ For the joint input-output distribution: $P_{\mathbf{i},j} = p_{\mathbf{i}} Tr\{\rho_{\mathbf{i}} A_j\}$, we denote the marginal output probability by: $p_j = \sum_{\mathbf{i}} p_{\mathbf{i}j}$. We define the joint probability between the output state and the first input bit (the 'WW bit') as, $p_{i_{\mathrm{ww}},j} \equiv \sum_{i_{\mathrm{wp}}} p_{(i_{\mathrm{ww}},i_{\mathrm{wp}})j}$ and similar for $p_{i_2,j}$. We define conditional probabilities such as $p_{i_{\mathrm{ww}}|j}$ and $p_{i_{\mathrm{ww}}|j_2}$ as well. We use

$$\mathbf{r}_{\mathbf{i}} = \mathbf{r}_{i_{\mathrm{ww}},i_{\mathrm{wp}}} = (d_0, i_{\mathrm{wp}} d_{\mathrm{wp}}, i_{\mathrm{ww}} d_{\mathrm{ww}}) \qquad (20)$$

for the Bloch representation of the $\mathbf{i}$-th input state, and $\mathbf{R}_j = (x_j, y_j, z_j)$ for that of A_j (the operator corresponding j-th measurement outcome), which is possible due to Lemma 1 above. In this notation we have:

$$\begin{aligned} p_{\mathbf{i}j} &= \frac{\mu_j}{8}(1 + d_0 x_j + i_{\mathrm{wp}} y_j d_{\mathrm{wp}} + i_{\mathrm{ww}} z_j d_{\mathrm{ww}}), && (21)\\ p_{i_{\mathrm{ww}}j} &= \frac{\mu_j}{4}(1 + d_0 x_j + i_{\mathrm{ww}} z_j d_{\mathrm{ww}}). && (22)\end{aligned}$$

Given that an output j has occurred, the most probable value, i_{ww}, of b_{ww} is that for which $p_{i_{\mathrm{ww}},j} > p_{-i_{\mathrm{ww}},j}$ The special case when these joint probabilities are equal needs to be treated slightly differently, but the results for the generic state can be shown to apply to it by continuity, and we shall not treat it here. From this follows that the total probability for correct WW inference is

$$P_{\mathrm{WW}} = \sum_j \max\{p_{i_{\mathrm{ww}},j}, p_{-i_{\mathrm{ww}},j}\}. \qquad (23)$$

By Eq. (21),

$$\max\{p_{i_{\mathrm{ww}},j}, p_{-i_{\mathrm{ww}},j}\} = \frac{\mu_j}{4}(1 + d_0 x_j + |z_j|\, d_{\mathrm{ww}}), \qquad (24)$$

therefore

$$P_{\mathrm{WW}} = \frac{1}{2}\left(1 + \sum \frac{\mu_j}{2}|z_j|\, d_{\mathrm{ww}}\right), \qquad (25)$$

where we have used $\frac{1}{2}\sum\mu_j = 1, \sum x_j = 0$. Similarly,

$$P_{\mathrm{WP}} = \frac{1}{2}\left(1 + \sum \frac{\mu_j}{2}|y_j|\, d_{\mathrm{wp}}\right). \qquad (26)$$

Combining these two equations, we have

$$\begin{aligned} &\left(\frac{2P_{\mathrm{WW}}-1}{d_{\mathrm{ww}}}\right)^2 + \left(\frac{2P_{\mathrm{WP}}-1}{d_{\mathrm{wp}}}\right)^2 = \\ &\left(\sum \frac{\mu_j}{2}|z_j|\right)^2 + \left(\sum \frac{\mu_j}{2}|y_j|\right)^2 \leq \\ &\sum \frac{\mu_j}{2}\left(z_j^2 + y_j^2\right) \leq \sum \frac{\mu_j}{2} = 1. \end{aligned} \qquad (27)$$

The last inequality follows from Jensen's inequality (for the concave function $f(x) = x^2$). ■

Let us now characterize the *Pareto-optimal POVMs.* The inequality becomes an equality iff $|z_j| = const. \equiv z_0, |y_j| = const. \equiv y_0, z_0^2 + y_0^2 = 1$. Clearly, we are free to choose $y_0 \in [0,1]$, and let $z_0 = \sqrt{1-y_0^2}$. Then the Bloch vectors $\mathbf{R}_j$ can take the values $(0, \pm y_0, \pm z_0)$, with corresponding weights $\mu_{\pm,\pm}$. The condition $\sum \mu_j \mathbf{R}_j = 0$ implies $\mu_{++} = \mu_{--}$, $\mu_{+-} = \mu_{-+}$ while $\sum \frac{1}{2}\mu_j = 1$ implies $\mu_{++} + \mu_{+-} = 1$. Therefore, the optimal POVMs are represented by rectangles residing on a great circle in the $y-z$ plane, with sides parallel to the rectangle of input states, and weights satisfying:

$$\mu_{++} = \mu_{--} = \mu, \;\; \mu_{+-} = \mu_{-+} = 1-\mu; \;\; \mu \in [0,1]. \qquad (28)$$

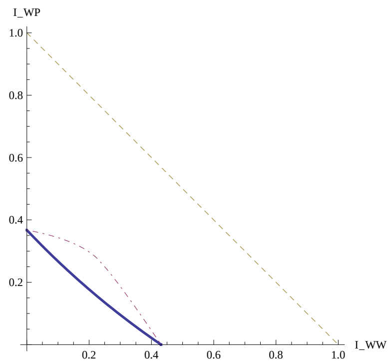


Figure 4: Comparison of the exact tradeoff given by Eq. (19) (solid line) with the one given by Eq.(30) (dashed); the dot-dashed line shows $I_{\text{in}-\text{out}} - I_{\text{WW}}$: the gap between it and the solid line is the contribution of $I_{\text{WW:WP}}$ to I_{in-out}. All of these relations generalize naturally to the three dimensional case ($I_{\text{WW}} + I_{\text{WP}} + I_{\text{WM}}$).

For the optimal POVMs, Eqs. (25, 26) simplify to

$$P_{\text{WW}} = \frac{1}{2}\left(1 + z_0 d_{\text{ww}}\right), \quad P_{\text{WP}} = \frac{1}{2}\left(1 + y_0 d_{\text{wp}}\right). \tag{29}$$

Bounds on $I_{\text{WW}} + I_{\text{WP}}$

We note that by Holevo's theorem[8] $I_{\text{in}-\text{out}} \leq S(\sum p_i \rho_i) - \sum p_i S(\rho_i) = H_2(\frac{1+d_0}{2})$, where S denotes the Von-Neumann entropy, and ρ_i , p_i denote the ith initial state and its a priori probability, respectively. However, this bound is not tight in the present situation. Together with Eq. (11), this implies the following bound:

$$I_{\text{WW}} + I_{\text{WP}} \leq H_2(\frac{1+d_0}{2}), \tag{30}$$

where $d_0 = \sqrt{1 - (d_{\text{WW}}^2 + d_{\text{WP}}^2)}$. As mentioned there, this bound is *not* tight. However, as explained following Eq. (10), the $P_{\text{WW}} - P_{\text{WP}}$ Pareto-frontier is also the frontier for $I_{\text{WW}} - I_{\text{WP}}$, as the Is are monotonic functions of the Ps. Thus, Eq. (19) implicity gives the exact tradeoff between the Is.

WW-WP-WM experiments

The Von Neumann projective measurement depicted in Fig. 2c can roam on the entire ellipsoid, provided the input phase on the output BS be tunable, as well as its bias. Conversely, any output beam splitter is Pareto-optimal. The WW detector generalizes to a two detector scheme, with detectors with efficiencies E_1, E_2 (Fig. 2d), yielding: $\frac{P_{\text{WW}}-\frac{1}{2}}{d_{\text{ww}}} = E_1$, $\frac{P_{\text{WP}}-\frac{1}{2}}{d_{\text{wp}}} = \sqrt{1-E_1^2}E_2$, $\frac{P_{WM}-\frac{1}{2}}{d_{\text{wm}}} = \sqrt{(1-E_1^2)(1-E_2^2)}$.